Providing a Foundation for Power and Resiliency

James H. Laros III*
*Sandia National Laboratories
Email: jhlaros@sandia.gov

I. Introduction

Among the four major challenges enumerated in the *Exascale Computing Study*[1] are, power, stated to be "the most pervasive of the four", and resiliency. Power and resiliency are cross cutting issues that have the potential to impact hardware design and most, if not all, levels of the system and user level software stack. At the core of both power and resiliency is the necessity to obtain information. The type and genesis of this information is diverse. At times the necessary information can be obtained locally. Often, a global perspective or information from remote components is necessary.

Once obtained, this information can be used to modify application behavior or platform characteristics, for example. This necessitates the ability to react, in some way, to the information collected. What we have described is analogous to a type of feed-back loop. Information is obtained in real or near real time, depending on the type of information. The *consumer* of the information, operating system or user application for example, reacts to the information in a range of potential ways. This reaction, in turn, can affect subsequent information.

The implementation implications of this feedback loop are many. Increasing core counts, while providing another challenge for Exascale, provide an opportunity to address power and resiliency challenges. Vendors are increasingly using one or a few of the cores on a node to isolate operating system services. One motivation of this is reducing operating system noise. If we assume this trend will continue, we can leverage this currently underutilized resource to implement the power/resilience feed-back loop previously described.

II. Challenges

Traditionally, Reliability Availability and Serviceability (RAS) systems have been designed to operate largely out-of-band from the primary platform. While this concept has enhanced resiliency in many ways, it has constructed a de facto wall between an abundance of potentially useful information and potential consumers (system and application level services) that could benefit from the information (and provide potential feedback). For example, while memory errors are typically monitored this information is seldom transmitted to the user level application. If provided, information about failing memory, or other components, could allow an application to positively affect its own resiliency. There are numerous other examples of information that is currently being collected that could improve application resiliency. Even more examples can be enumerated where the application can direct resiliency by informing the platform of errors or faults that can be tolerated and others that cannot.

To affect power use, and subsequent energy efficiency, at Exascale both node and platform level information must be made available. Currently, there are few demonstrated examples of fine granularity power monitoring at scale[2], and none to our knowledge that have been implemented as a production capability. Certainly, this information is not currently available to the user application. Recently, chip manufactures are trending towards providing some level of power and energy monitoring on chip. High frequency external monitoring is also gaining ground as a necessary feature on next generation architectures. Regardless of the source of information, the driving requirement is that power and energy use must be monitored at component level granularity. This information must be made available not only to systems administrators via a segregated RAS system but to system services (node-level operating system) and user applications. Once made available, the operating system or user application must be able to affect power/energy use at the node or platform level, by tuning CPU frequency for example.

The common thread shared by both resiliency and power is the need for a link that not only enables exposure to important information but allows feedback from the information consumer (system or user level software). While increasing node core counts provides an additional challenge for Exascale, this trend provides an opportunity to rethink some of the historical tenants of how RAS systems are implemented, how they can more actively enhance resiliency and how power and energy use can be monitored and affected.

III. Proposal

We propose the following:

- Bridge the gap between existing RAS information and system and user level software on the platform. RAS data is critical to improve resiliency.

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

- Leverage the same mechanism to complete a feedback loop for the purposes of affecting and/or controlling resiliency.
- Create the ability to convey node application and platform level power and energy information to system and user level software.
- Complete the power and energy feedback look by providing a mechanism to affect power and energy use ranging from component to platform level. This could take the form of power capping or frequency scaling, for example.
- Develop an Application Programming Interface (API) that defines how to mine the available information and affect platform operational and performance characteristics.

The solution proposed to satisfy these Exascale challenges requires both hardware and software development and innovation. While we can expect that there will be sufficient node level resources (spare cores) to satisfy part of the necessary hardware infrastructure, hardware channels must exist that allow two way connectivity (information and control) from the core or cores to other board level components and the out-of-band RAS network infrastructure. Additionally, the instrumentation necessary to measure power (current and voltage) over time (energy) must be present for components, nodes, boards, racks or cabinets, depending on the underlying architecture and the granularity of control. The API must allow for reasonable diversity in measurement and control capability among different and evolving architectures.

The effort required to successfully implement this capability is significant. Portions of this effort require hardware capabilities that while not currently available in the required form, do currently exist as basic technologies. Given necessary hardware capabilities, this effort requires significant software design and development with great attention paid towards scale issues. While some of these challenges and usage models are specific to Exascale platforms, at a fundamental level this ability could provide great utility to a wider range of platforms. It is critically important that the development target is Exascale. Historically, we have seen time and time again that technologies developed for the enterprise and small scale fall far short of meeting requirements at large scale.

The development of the low level hardware interfaces would require direct communication between hardware architects and system software experts. The development of the API would require close collaboration between system and application level software experts. Both resiliency and power are cross cutting issues which simultaneously require a high level architectural view while considering the details of low level interfaces.

References

[1] K. Bergman, S. Borkar, D. Campbell, W. Carlson, W. Dally, M. Denneau, P. Franzon, W. Harrod, J. Hiller, S. Karp, S. Keckler, D. Klein, R. Lucas, M. Richards, A. Scarpelli, S. Scott, A. Snavely, T. Sterling, R. S. Williams, K. Yelick, K. Bergman, S. Borkar, D. Campbell, W. Carlson, W. Dally, M. Denneau, P. Franzon, W. Harrod, J. Hiller, S. Keckler, D. Klein, P. Kogge, R. S. Williams, and K. Yelick, "Exascale computing study: Technology challenges in achieving exascale systems peter kogge, editor& study lead," 2008.

[2] J. H. Laros III, K. T. Pedretti, S. M. Kelly, J. P. Vandyke, K. B. Ferreira, C. T. Vaughan, and M. Swan, "Topics on measuring real power usage on high performance computing platforms," in *International Conference on Cluster Computing*. IEEE, 2009.